CLAIMS

1. Apparatus for identifying animal species from their vocalizations, comprising:

a source of digital signal representative of at least one animal candidate vocalization;

a feature extractor that receives the digital signal, recognizes notes therein and extracts phrases including plural notes and that produces a parametric representation of the extracted phrases; and

a comparison engine that receives the parametric representation of at least one of the digital signal and the extracted phrases, and produces an output signal representing information about the animal candidate based on a likely match between the animal candidate vocalization and known animal vocalizations.

2. The apparatus as claimed in claim 1, wherein the feature extractor comprises:

a transformer connected to receive the digital signal and which produces a digital spectrogram representing power and frequency of the digital signal at each point in time.

3. The apparatus as claimed in claim 2, wherein the transformer comprises:

a Discrete Fourier Tansformer (DFT) having as an output signal a time series of frames comprising the digital spectrogram, each frame representing power and frequency data at a point in time.

4. The apparatus as claimed in claim 2, wherein the power is represented by a signal having a logarithmic scale.

5. The apparatus as claimed in claim 2, wherein the frequency is represented by a signal having a logarithmic scale.

6. The apparatus as claimed in claim 2, wherein the power is represented by a signal that has been normalized relative to a reference power scale.

7. The apparatus as claimed in claim 2, wherein the frequency is represented by a signal that has been normalized relative to a reference frequency scale.

8. The apparatus as claimed in claim 1, wherein the feature extractor further comprises a discrete cosine transform (DCT) transformer receiving the digital signal and producing a signal representing plural coefficients defining the parametric representation of the extracted phrases.

9. The apparatus as claimed in claim 1, wherein the feature extractor further comprises:

a transformer connected to receive the digital signal and which produces a signal defining a parametric representation of each note.

10. The apparatus as claimed in claim 9, wherein the transformer is a discrete cosine transform (DCT) transformer.

11. The apparatus as claimed in claim 9, wherein the feature extractor further comprises:

a time normalizer operative upon each note recognized in the digital signal before the transformer receives the digital signal.

12. The apparatus as claimed in claim 9, wherein the comparison engine further comprises:

a cluster recognizer that groups notes into clusters according to similar parametric representations.

13. The apparatus as claimed in claim 12, wherein the cluster recognizer performs K-Means.

14. The apparatus as claimed in claim 12, wherein the cluster recognizer is a self-organizing map (SOM).

15. The apparatus as claimed in claim 12, wherein the cluster recognizer performs Linde-Buzo-Gray.

16. The apparatus as claimed in claim 1, wherein the comparison engine further comprises:

a neural network trained to recognize likely matches between the animal candidate vocalization and the known animal vocalizations.

17. The apparatus as claimed in claim 16, wherein the neural network further comprises:

plural layers of processing elements arranged between an input of the comparison engine and an output of the comparison engine, including a Kohonen self-organizing map (SOM) layer.

18. The apparatus as claimed in claim 16, wherein the neural network further comprises:

plural layers of processing elements arranged between an input of the comparison engine and an output of the comparison engine, including a Grossberg layer.

19. The apparatus as claimed in claim 1, wherein the comparison engine further comprises:

a set of hidden Markov models (HMMs) excited by the parametric representation received, each HMM defined by a plurality of states.

20. The apparatus as claimed in claim 19, wherein at least one of the plurality of states comprises:

a data structure holding values defining a probability density function defining the likelihood of producing an observation.

21. The apparatus as claimed in claim 20, wherein the probability density function is a multi-variate Gaussian mixture.

22. The apparatus as claimed in claim 21, wherein the multi-variate Gaussian mixture is defined by a fixed co-variance matrix.

23. The apparatus as claimed in claim 19, wherein an HMM of the set of HMMs produces an observation corresponding to a bird species.

24. The apparatus as claimed in claim 19, wherein an HMM corresponding to a set of training data representing at least one vocalization comprises:

a first set of states representing a first cluster of time-normalized notes, classified according to similar parametric representations; and

a second set of states representing a second cluster of time-normalized notes, classified according to similar parametric representations different from those of the first cluster of time-normalized notes.

25. The apparatus as claimed in claim 24, wherein the HMM further comprises:

a state corresponding to a gap between a note of the first cluster and a note of the second cluster.

26. The apparatus as claimed in claim 24, wherein the set of training data includes coefficients from a discrete cosine transform (DCT) performed on a vocalization signal.

27. The apparatus as claimed in claim 24, wherein the first cluster comprises classification vectors clustered together using a K-Means process.

28. The apparatus as claimed in claim 24, wherein the first cluster comprises classification vectors clustered together using a self-organizing map (SOM).

29. The apparatus as claimed in claim 24, wherein the first cluster comprises classification vectors clustered together using Linde-Buzo-Gray.

30. The apparatus as claimed in claim 1, further comprising a database of known bird songs.

31. The apparatus as claimed in claim 30, wherein the database comprises: a data structure holding values in a memory of weights for a neural network.

32. The apparatus as claimed in claim 30, wherein the database comprises: a data structure holding values in a memory of parameters for a hidden Markov model (HMM).

33. The apparatus as claimed in claim 30, wherein the database comprises: a data structure holding records in a memory corresponding to the known bird songs specific to at least one of a region, a habitat, and a season.

34. The apparatus as claimed in claim 30, wherein the database of known bird songs is stored in a replaceable memory, such that the database of known bird songs can be modified by replacing the replaceable memory with a replaceable memory holding the modified database.

35. The apparatus as claimed in claim 30, wherein the database of known bird songs is stored in a modifiable memory.

36. The apparatus as claimed in claim 35, wherein the apparatus includes a port through which modifications to the database of known bird songs can be uploaded.

37. The apparatus as claimed in claim 36, wherein the port is wireless.

38. The apparatus as claimed in claim 1, further comprising:

a digital filter interposed between the source of a digital signal and the signal analyzer and classifier.

39. The apparatus as claimed in claim 1, wherein the source further comprises:

a microphone.

40. The apparatus as claimed in claim 39, wherein the source further comprises:

an analog-to-digital converter connected to receive an analog signal from the microphone an to produce the digital signal.

41. The apparatus as claimed in claim 39, wherein the microphone further comprises:

a shotgun microphone.

42. The apparatus as claimed in claim 39, wherein the microphone further comprises:

a parabolic microphone.

43. The apparatus as claimed in claim 39, wherein the microphone further comprises:

an omnidirectional microphone.

44. The apparatus as claimed in claim 39, wherein the microphone further comprises:

an array of microphones.

45. The apparatus as claimed in claim 44, wherein the array of microphones is made directional by use of beam-forming techniques.

46. The apparatus as claimed in claim 1, wherein the source further comprises:

an analog signal input; and

an analog-to-digital converter connected to receive a signal from the analog input, and producing the digital input signal.

47. The apparatus as claimed in claim 1, wherein a time from the signal transformer receiving the digital signal to the comparison engine producing the output signal is real-time.

48. A computer-implemented method of identifying animal species, comprising:

obtaining a digital signal representing a vocalization by a candidate animal;

transforming the digital signal into a parametric representation thereof;

extracting from the parametric representation a sequence of notes defining a phrase;

comparing the phrase to phrases known to be produced by a plurality of possible animal species; and

identifying a most likely match for the vocalization by the candidate animal based upon the comparison.

49. The method of claim 48, wherein comparing further comprises:

applying a portion of the parametric representation defining the phrase to plural Hidden Markov Models defining phrases known to be produced by a plurality of possible animal species; and

computing a probability that one of the plurality of possible animal species produced the vocalization by the candidate animal.